Original Paper

Title:

Spelling errors and 'shouting' capitalization lead to additive penalties to trustworthiness of online health information: Randomized Experiment with Laypersons

Authors

Harry J Witchel [1,*], Georgina A Thompson [1], Christopher I. Jones [1], Carina E I Westling [2,†], Juan Romero [3], Alessia Nicotra [3], Bruno Maag [3], Hugo D Critchley [1]

Addresses

1 Department of Neuroscience, Brighton and Sussex Medical School, Brighton BN1 9RY, UK
2 School of Media, Film and Music, University of Sussex, Brighton BN1 9RH, UK
3 Dalton Maag Ltd., London SW9 9SP, UK

Georgina Thompson ORCID: 0000-0001-6777-5730
Carina Westling ORCID: 0000-0002-6549-0974
Alessia Nicotra ORCID: 0000-0001-7775-0166
Bruno Maag ORCID: 0000-0002-2841-8439
Hugo Critchley ORCID: 0000-0002-2445-9284
Juan Romero ORCID: 0000-0001-5972-4001
Christopher Jones ORCID: 0000-0001-7065-1157
* Author for correspondence: h.witchel@bsms.ac.uk, ORCID 0000-0001-8404-3494
† Current address: Faculty of Media and Communication, Bournemouth University, Bournemouth BH12 5BB, UK

Short Title

"Trustworthiness spelling capitalization"

Keywords

trust, trustworthiness, credibility, incivility, information quality, grammatical errors, writing mechanics, capitalisation

Abstract

BACKGROUND. The written format and literacy competence of screen-based texts can interfere with the perceived trustworthiness of health information in online forums, independent of the semantic content. Unlike in professional content, the format in unmoderated forums can regularly hint at 'incivility', perceived as deliberate rudeness or casual disregard towards the reader, e.g. through spelling errors and unnecessary emphatic capitalization of whole words (online 'shouting').

OBJECTIVE. To quantify the comparative effects of spelling errors and inappropriate capitalization on ratings of trustworthiness independently of lay insight, and to determine whether these changes act either synergistically or additively on the ratings.

METHODS. In online experiments, 301 UK-recruited participants rated thirty-six randomised stimulus paragraphs in the format of information from an unmoderated health forum (about multiple sclerosis) for trustworthiness using a semantic differential slider. Nine control paragraphs were compared to error-containing paragraphs including 5 instances of misspelling, 5 instances of inappropriate capitalization (‘shouting’), or a combination of the two. Data were analysed in a Linear Mixed Effects model.

RESULTS. The mean trustworthiness ratings of the control paragraphs ranged from 32.59 to 62.31 (rating scale 0-100). Compared to the control paragraphs, paragraphs containing only misspellings were rated as being 8.86 points less trustworthy, inappropriate capitalization was 6.41 less, and the combination of misspelling and capitalization was 14.33 less. Misspelling and inappropriate capitalization show an additive effect ($P < .05$ for all).

CONCLUSIONS. Distinct indicators of incivility independently and additively penalize perceived trustworthiness of online text independently of lay insight, eliciting a medium effect size.

Introduction

Trustworthiness of online health information: background, context and importance

As of 2019, 90% of all British adults use the internet at least weekly [1 Office for National Statistics 2019], and as patients they often search online for health information in order to solve their medical problems [2 Chen 2018]; furthermore, they are likely to be influenced by the online information, including changing their healthcare decisions and their frequency of ambulatory care visits [3 Hsieh 2016]. In a survey of 2,200 adults with chronic health conditions in the USA who were active social media users, 57% used a health-condition specific website (e.g. specialising in multiple sclerosis or rheumatoid arthritis) on a monthly basis, and 5% used such sites daily; half of the patients surveyed had asked a health-related question to others online within the previous 6 months, and 87% of those were seeking responses from other patients with the health condition [4 Health Union 2016].

In response to this explosion of unvetted potential sources providing online healthcare information that is acted upon, researchers, experts and medical professionals have repeatedly expressed concern about the inaccuracy of online information and the limited ability of lay consumers to adequately assess its validity [2 Chen 2018, 5 Diviani 2015, 6 Sun 2019]. In particular, it has long been known that when lay users determine whether to use and trust online healthcare information, they are strongly influenced by non-medical criteria that experts do not use [7 Fogg 2003b, 8 Lederman 2014, 9 Stanford 2002, 10 Brand 2017, 11 Sillence 2006]. Broadly, while academics favour a checklist approach of transparency criteria [12 England 2004], the approach of non-experts appears to be more variable, and situation-dependent, and it prioritises factors such as ease of understanding and the attractiveness of graphic design [9 Stanford 2002]. More generally, in research on the factors that influence judgments of trust, all aspects of trustworthiness can be relational and depend on the type of person (or group being studied); a major relational factors include accessibility, both cognitive as well as physical [13 Woudstra 2016], and correctly accommodating language to the intended audience [14 Zimmerman 2018].

Understanding lay assessment of trustworthiness of online health information is important because false online information presented to the general public, if believed, has the potential to undermine correct medical advice [15 Kata 2012], to elicit unhealthy behaviour [16 Borzekowski 2010] as well as influence socio-political discourses on healthcare and other topics [17 Flynn 2017, 18 Picard 2017, 19 Hobbs 2004]. Sbaffi and Rowley's [20 Sbaffi

2017] recent review of how lay people assess the trustworthiness and credibility of online health information concluded that thus far much less research has been performed to understand what interferes with trust (we describe these as penalties to trustworthiness [21 Albuja 2018]), compared to what causes trust.

Trust and credibility are closely related to one another and to information quality; although there remains disagreement among researchers as to the exact definitions of these terms. In general the definitions emphasise likelihood of information use, believability, reliability and dependability [22 Rieh 2007, 20 Sbaffi 2017]. Correspondingly, there is little evidence to suggest that the general population reliably make fine conceptual distinctions between trustworthiness and credibility of information.

Online health support often is divided into seeking information or emotional reassurance, and can be gender-specific (e.g. prostate cancer versus breast cancer) and person-specific [23 Gray 1996, 24 Seale 2016]. Trustworthiness remains pertinent to online emotional reassurance and sharing, as shown by the many occurrences of large-scale hoaxes designed to manipulate emotions [8 Lederman 2014]. For example, between 2010 and 2011 a Macmillan cancer forum was inundated with posts in response to an elaborate hoax by a purported mother about her six-year-old daughter struggling with cancer. Upon the exposure of the hoax (perpetrated by a lonely sixteen-year-old girl), many users of the forum – who had formed close online relationships with the supposed mother -- refused to believe it had all been a hoax [25 BBC 2012]. In light of the range of uses that online health information has for the general public, we contend that it is important to build knowledge of the factors that influence how trust and its absence are formed online beyond source analysis and fact checking. As this study shows, both linguistic and meta-linguistic factors affect how trustworthiness is instinctively rated.

Theoretical Underpinnings for Factors Influencing Trustworthiness

The range of elements that influence online trustworthiness includes issues related to the source (e.g. author identification and the absence of advertising), issues related to the content (e.g. a date stamp and inclusion of medical evidence), issues relating to design and engagement (e.g. inclusion of images) and issues affecting all of the above (e.g. the absence of typographical errors) [20 Sbaffi 2017, 26 Metzger 2007]. For some time, credibility has been subdivided into aspects such as source credibility, message credibility and media credibility – which strongly influence each other [22 Rieh 2007]. All such aspects of

credibility include accessibility, which can be relational and cognitive as well as physical [13 Woudstra 2016]. More recently, Sun et al. [6 Sun 2019] have divided the elements that influence consumer evaluation of online health information into 25 *criteria* and 165 *quality indicators*; criteria are rules that reflect notions of value and worth (e.g. expertise and objectivity), while quality indicators are properties of information objects to which criteria are applied to form judgements (e.g. the owner of the website and inclusion of statistics). In line with Diviani et al. [5 diviani 2015], Sun et al. [6 Sun 2019] have suggested that indicators can be positive or negative (in terms of trustworthiness), and that consumers' perceived online health information quality could conceivably be measured by a small set of core dimensions (i.e. a few groups of criteria might incorporate many of the quality indicators that explain most of the trustworthiness judgments).

Thus far, Lederman et al. [8 Lederman 2014] have proposed a five-category model that highlights verification processes via the comparison of different web sites or other online statements. This model includes argument quality, source credibility, source literacy competence and crowd consensus. An extended six-category model [26 Metzger 2007] proposes that the lay reader may assess some or all of the following: reputation, endorsement, consistency with other sources, self-confirmation (agreement with the reader's own opinions), persuasive intent, and expectancy violation. For the purposes of our research, we have adopted Lederman et al.'s [8 Lederman 2014] model of credibility; however, the choice among these models for our current research is moot because all the models [20 Sbaffi 2017, 26 Metzger 2007, 5 diviani 2015, 6 Sun 2019, 22 Rieh 2007] are concordant with the idea that spelling errors will detract from judgments of trustworthiness (i.e. they are negative quality indicators).

Incivility, Literacy Competence and Errors of Writing Mechanics

Although institutionally produced web sites and curated online health content (cf. [6 Sun 2019]) will usually be both grammatically correct and civil, the responses by the general public may be uncivil [27 Kornfield 2015]. Inappropriate capitalization (including inappropriate capitalization of entire words, sometimes called online 'shouting') and misspelling are examples of errors in literacy competence [8 Lederman 2014] and/or writing mechanics [28 Ketron 2017]; writing mechanics is defined as elements of a language that only manifest when communication is in written form. Both inappropriate capitalization and misspelling have been highlighted by qualitative investigations as explicit criteria used by lay readers in judgements of online credibility [8 Lederman 2014, 29 Schindler 2012]. The

rationales given to explain why these two error types undermine trustworthiness are that either (1) the errors imply a lack of intelligence (expertise, ability, authority) [30 Figueredo 2005, 8 Lederman 2014], or (2) that they suggest a lack of motivation (objectivity, attention to detail, conscientiousness) to be trustworthy [31 Morin-Lessard 2017, 32 Vignovic 2010]. The term incivility is used to describe this latter lack of motivation or effort to make statements that are compliant with rules of communication. Incivility implies a lack of respect for the reader, the platform, and/or the rules of social exchange [33 Sobieraj 2011, 34 Gervais 2015], and it is fundamentally relational. Quantitative research on incivility in the mainstream world wide web demonstrates that civil statements are rated as more trustworthy and influential than uncivil ones [35 Graf 2017, 36 Thorson 2010].

Integration versus Heuristics: Lay Judgments Based on Multiple Cues

The reader who must judge unvetted online health information is faced with a wealth of cues that indicate the degree of trustworthiness, and those cues may have contradictory effects (e.g. a cogent message that is misspelled). There are three broad theories for how individuals (both lay and expert) make judgments based on multiple cues. The rational approaches are represented by the Information Integration model [37 Anderson 2013], in which an individual accounts for all the different pieces of information by a complex (but often subconscious) mathematical process that is usually based on addition, multiplication or averaging; extensive observations of integration in judgments occur across cultures and individuals. This is the process that Sun et al. [6 Sun 2019] allude to when they propose that the consumer integrates the relevant trustworthiness criteria and quality indicators in a "complex cost-benefit analysis". In 2003 Fogg's Prominence Interpretation Theory proposed explicit mathematical relationships for how to predict the effects of multiple factors on credibility, but the theory never detailed how to measure the relevant quantities independently [38 Fogg 2003a]. Computational models typically assume that the elements of incivility (e.g. inappropriate capitalization) act together either additively or non-linearly [39 Wanas 2008, 40 Castillo 2013, 41 Weerkamp 2012], although hypothesis-led proof for this assumption is minimal. The elucidation of this integration is only just starting in the literature [42 Johnson 2015], and the relative importance of each indicator in this intuitive cost-benefit analysis remains unknown; the relative values for each cue may be elucidated empirically by statistical methods such as regression, but it is unlikely that these values would be explicit or transparent in the minds of most lay decision makers. This is an area that needs to be further researched.

An alternative theory to costs-benefit analysis for how judgments are made based on multiple cues is the process-level cognitive perspective [43 Pachur 2013]; this has been made famous by the heuristics and biases research program from behavioural economics [44 Tversky 1974]. Heuristics are rapid cognitive strategies (either explicit or subconscious) formulated as rapid bounded rational decision systems for multiple cues that can be more transparent than complex cost-benefit analyses, for example, hierarchical lexicographic decision models [45 Gigerenzer 1996]. In a take-the-best judgment [46 Gigerenzer 2011] first a single cue (the most important one) is searched for in the environment and considered independently of all others, and if a tie or no clear result occurs, then the second most important cue is sought and decided upon, and so on. For example, when deciding whether you have the right of way when driving your car through an intersection, first you follow the signal of any policemen present, and only if there are no policemen present do you seek and consider a traffic light (including a temporary traffic light for construction), and only if there are no traffic lights present do you then consider static road markings and the positions of the other cars.

In biased heuristics, only a limited subset of information (often only one cue) is used to make a fast and ecological judgment outside of conscious awareness [43 Pachur 2013]; when biased, these heuristics are used to support preferred or preconceived outcomes. With biased heuristics, the prioritization of the cue, and even the cue's basic validity for the judgment being made, is dubious. Such biased heuristics have been used to explain seemingly irrational preferences that individuals make in situations involving slot machines and organ donation [47 Bennis 2012]. Examples of biased heuristic processes include representativeness -- where some cues are weighted disproportionately compared to their real representativeness -- and availability -- where a cue that is easily recalled (such as occurs with recency and news) determines the judgment. None of these judgment models so far proposed have explicitly assessed specific issues within literacy competence.

Determining the Criteria and Weighting of Judgments Based on Multiple Cues

A key feature of heuristics is that they are typically made subconsciously, and that *post hoc* explanations for such choices are often self-serving justifications or rationalizations [48 Oreg 2009]. That is, in the case of heuristics, the decision maker does not have privileged information on how the judgment was made, and, furthermore, the decision maker can be wrong about themselves [49 Koriat 2015]. For example, university students have been shown to greatly over-estimate how much they *actually learned* from excellent lecturers (conflating it with how much they *feel* they learned, which is discounted by effort and

exertion) [50 Deslauriers 2019]. This creates potential issues for researchers, in which insight-based techniques (both qualitative interviews as well as quantitative questionnaires) can lead to judgments and explanations of causes that are inaccurate because of demand characteristics or social desirability [51 Orne 2009; 52 Krefting 1991]. It has long been known that when evaluating credibility (e.g. website privacy policy), the importance of factors that lay individuals say they would use to make their judgment diverge from the factors they are observed to use [38 Fogg 2003a]. Recent sophisticated measurements suggest that, although multiple factors can interact when quantifying perceived credibility (of information for an online health forum), these interactions do not support Petty and Cacioppo's Elaboration Likelihood Model [53 Petty 1984]. Furthermore, there is a gap in the literature for any data showing how spelling errors interact with other writing mechanics violations that might also affect trustworthiness [54 Sauls 2018]. This suggests that lay insight into the influences on their perceptions of trustworthiness are imperfect, and that research on trustworthiness should be supplemented by approaches that do not rely upon such insight.

In order to avoid insight bias from our lay participants, the approach of this study is to compare and contrast **marginal differences** in penalties to trustworthiness elicited by different combinations of literacy or writing mechanics violations. Our experiments on changes in marginal trustworthiness were specifically organised so that the participating healthy volunteers were unaware that the experiment tested the effects of capitalization or misspelling *per se* (with full institutional ethical approval). We did not explicitly ask lay individuals for their beliefs regarding how their process for judging message credibility incorporates quality indicators relating to source credibility and media credibility. That is, instead of asking directly, "How much less would you trust a web post that is misspelled?", we simply presented the participants with some posts that had misspellings and asked the same question as usual, "How trustworthy do you find this information?" while subtly varying misspelling and capitalization (i.e. ‘shouting text’, not acronyms or beginnings for sentences).

Research Questions

RQ1: Do errors in writing mechanics and incivility lead to marginal changes (i.e. without signposting) in judgments for message trustworthiness?

RQ2: Are the marginal effects of incivility, such as with inappropriate capitalization, on trustworthiness judgments quantitatively comparable to the known effects of writing mechanics such as spelling errors?

RQ3: How, if at all, do the effects of writing mechanics (e.g. spelling errors) and incivility (e.g. inappropriate capitalization) integrate? Is there a ceiling effect or a binary effect in which once a message source is judged as incompetent, no further trustworthiness penalty is added to the judgment (similar to a take-the-best heuristic) [46 Gigernezer 2011]? Or is there some additive, multiplicative, synergistic or otherwise integrative function that increases the penalty on the judgment to a new higher level when both cues are present [37 Anderson 2013]?

Methods

The project was approved by our local ethics committee (Brighton and Sussex Medical School's Research Governance and Ethics Committee (RGEC, University of Sussex), approval 16/044/WIT). All experiments were performed according to the Declaration of Helsinki. All individuals provided informed consent via a welcome page in each online study. Participants were recruited via the prolific.co website. we specified that it should focus on UK-based members of the public. Participants consented to participate with the understanding that the research concerned "responses to online text"; none of the advertising, web URLs or experimental information to the participants mentioned that the experiment was related to graphics, formatting, incivility, spelling, etc. This feature of the advertising was approved by our ethics committee, not least because the paragraphs were not considered misleading or potentially emotionally adverse.

Study Design Process

The study design was a confirmatory, cross-sectional experiment with a balanced incomplete block design; it was a randomized experiment with lay participants, each of whom experienced only a limited number of the possible options. The total number of experimental excerpts that were tested in the entire cohort was 36; there were 9 excerpts, each in 4 possible versions: no errors, inappropriate capitalization only, misspelling only, or the combination of both errors. The stimulus texts (and the versions) are shown in the supplementary methods YU9WV2. There were two additional paragraphs that always appeared as the first two paragraph stimuli, which were training stimuli. The training stimuli were not labelled as being different from other stimuli in any way, and were never included in the statistical

analyses. The only purpose of the training stimuli was to allow participants to familiarise themselves with the rating task, and with the range of the trustworthiness scale: Training stimulus 1 was quite believable (mean trustworthiness rating = 54.33 ± 23.39, mean ± sd, n = 301), and training stimulus 2 was less plausible (mean trustworthiness = 40.20 ± 25.80, n = 301, P < .001, paired t test).

Each participant experienced and rated only 9 (of the possible 36) experimental excerpts, and those 9 included exactly one version of each excerpt, and among those 9, that participant would experience a mix of different error types (Figure 4RV2zC). For example, participant number 001 experienced the capitalization-only version of excerpt E07, then the "both errors" version (capitalization and misspelling) version of excerpt E02, etc. This double randomization prevented any participant from seeing the same excerpt twice.

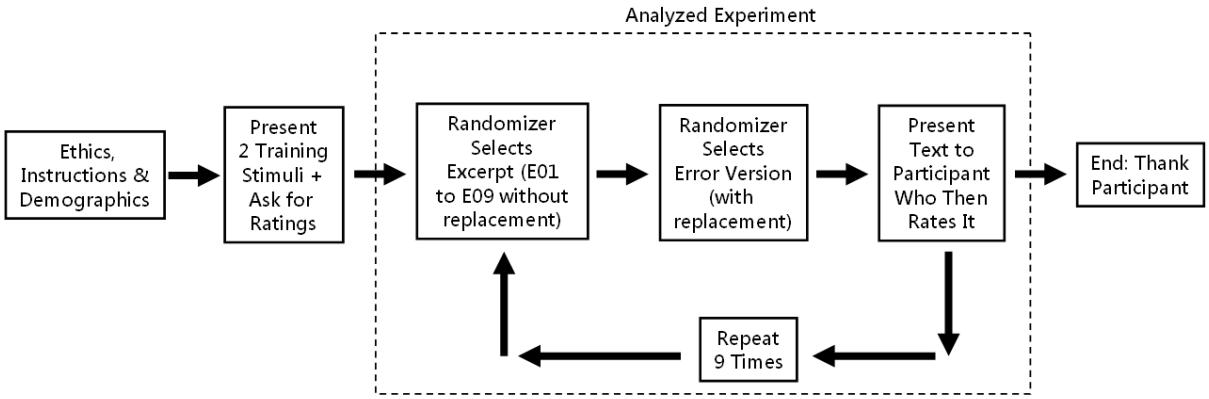

Figure 4RV2zC. Experiment design: what each participant experiences.

The goal was to present to participants a set of coherent paragraphs of similar lengths (70-100 words) on the topic of multiple sclerosis, each excerpt being a coherent answer to a question. The rationale for presenting excerpts in a question-then-answer format was that it was possible to ask whether readers trusted the advice enough to believe it or act upon it. A range of comments were found in the public domain (see Table HFH4RT), although the texts often had to be edited substantially to fit within the word count or to avoid explicitly endorsing commercial products (see Supplementary Methods PL9PdX for original and presented texts). Initially the paragraphs and the presentation system were tested online by a small group of testers, who then provided verbal feedback on the test to the experimental team. After that a cohort of 40 participants was recruited online to test the paragraphs and demonstrate that the paragraphs elicited similar standard deviations in ratings (19 to 28 units out of 100) and elicited a wide range of mean trustworthiness ratings (from 20 to 70). These excerpts were to

be presented either as they were ("no errors" or the negative control) or in one of the three error versions stated above.

Code	Brief Topic Description	Website	Length
L01	Numerous artificial sweeteners	blogspot.com	78
L02	Hoax about artificial sweeteners	quora.com	81
E01	Triggers of the immune system	healingchronicles.com	81
E02	Programmer's intelligence	dailystrength.org	89
E03	Epstein Barr Virus (EBV)	medicaldaily.com	92
E04	Avonex patient	quora.com	74
E05	Up there in risk	quora.com	96
E06	Mental exercises	dailystrength.org	85
E07	Vitamin D	quora.com	100
E08	Small risk of PML	my-ms.org	90
E09	Half of all people	ms.pitt.edu	71

Table HFH4RT. Sources of excerpts on multiple sclerosis.

Text Interventions

For the error versions of the texts (inappropriate capitalization or misspelling), we wanted to include 5 of the relevant errors per excerpt (10 errors total for the combination of both errors version), with the errors spread throughout the excerpt (rather than bunched together). In paragraphs where inappropriate capitalization was required, there would be five "sets" of words/phrases. Normally a set was one or two words, although one of the five sets had to be a 4-word series. The priorities for selecting words to capitalize were (in this order):

1. adverbs (especially those suggesting extremity such as "very" or "never")
2. judgements ("rubbish", "hopeless", "horror")
3. strong emotions ("worried", "angry")
4. words implying danger ("fatal", "death")
5. amounts ("all", "every", "ten")
6. adjectives (rather than nouns)
7. action verbs (especially gerunds)
8. conjunctions ("and")

To verify that each word that we capitalized was naturally capitalized on the web, we analyzed words that were capitalized online on Twitter. We used the Claritin corpus, which is a crowd sourced data set of all the Twitter tweets that contained the word "claritin" in the month of October 2012 [55 Oleson 2013]. This corpus has some 4,900 tweets, and we used Matlab to find all the words that were in all capitals (which did not have a hashtag or an at-

sign in them); this resulted in a list of 343 capitalized words (see Supplementary Methods TY3W1A), many of which were short words, acronyms and internet memes. From this list of words spontaneously capitalized on the web, we selected words in our excerpts to capitalize.

The rationale for how we selected words to misspell was that misspellings should be quite noticeable, and that the meaning of the words should remain clear to the reader even when misspelled. We avoided homonyms and words that looked plausibly English when misspelled. To make sure that misspelled words were noticeable, short words were preferred, or we placed the misspellings in the first syllable of a multi-syllable word. In addition, one of the misspelled words had to be in the first five words of a paragraph. The misspelled word had to be completely understandable (in the absence of other words or context) even when misspelt. Thus, a misspelled word with missing or added letters should be pronounceable in English (e.g. "yu" plainly means "you"). The types of misspellings were:

1. swap one letter for another letter that is next to it on a qwerty keyboard
2. double a consonant ("esstimate")
3. double a vowel, or add an extra vowel ("syystem")
4. leave out a vowel in the centre of the word (leaving only consonants: "xposure")

To verify that each word that we misspelled was naturally misspelled on the web, we searched for the misspelled word along with the words "health" and "forum"; if we could not find at least two examples of a misspelling on online health forums, we did not use it. A complete listing of the misspellings and where we found them online is in supplementary methods NC9W22.

Study Delivery

The study was presented to participants using the Qualtrics portal, which allows for a wide range of question types and keeps track of answers and total response time. A full description of the survey in the CHERRIES format [56 Eysenbach] is included as part of the supplementary methods FHKhq3 for this paper. The online study welcome page explained in brief what the study was about and what it entailed (estimated 8 minutes participation time, including reading the ethics and filling in demographics), the ethics of the study (include the ability to withdraw at any time), and a brief complaints procedure. The ethics page explicitly excluded participants under 18 years of age, or those from vulnerable populations. A pointer to a full length Participant Information Sheet (3 pages) was shown; the welcome/ethics page had an "I agree" button at the bottom. After the welcome page, participants filled in a brief

multiple choice demographics page, which included questions on gender, age, field of work (e.g. healthcare, agriculture, retired), and familiarity with English language/Roman alphabet. All demographics questions included an option for "rather not say". After the demographics page, participants saw the instructions page, and then were launched into the experimental ratings pages.

Each ratings page consisted of a short excerpt of text (which was randomized as to whether or not it had the spelling or capitalization errors), followed by a horizontal slider for rating how "trustworthy" the participant found the statement to be; the slider had anchors of "completely untrustworthy" (left) and "completely trustworthy" (right, see figure gJU7EE). Although the data collection was numerical (0-100, left to right), there were no numerical cues or anchors seen by the participants. The instructions to the participants for the trustworthiness ratings were, "If you find something trustworthy, you would be prepared to act upon it; an untrustworthy statement you would ignore, and a rating in the middle represents information where you would want more proof or confirmation that it is correct". As explained in the instructions to participants, each stimulus excerpt was written as if it was an answer to a question written on an unmoderated health forum, with a specific focus on multiple sclerosis. Multiple sclerosis was chosen as a topic because the information was obviously important, but healthy participants would be unfamiliar with the veracity of each statement; thus, we predicted that the trustworthiness ratings would be more susceptible to non-verbal cues. The questions were:

1. Is multiple sclerosis preventable?
2. How risky is Tecfidera as a treatment for multiple sclerosis?
3. Does multiple sclerosis decrease intelligence/IQ?

The nominal responses to these questions were the experimental stimulus excerpts being rated.

Figure gJU7EE. Unnumbered horizontal slider for trustworthiness ratings.

Study Design, Analysis and Statistics

To detect a difference in rating scores between two different stimulus paragraphs with 80% power at the 5% significance level, assuming the standard deviation is 30 and the difference in scores is 15 (equivalent to a medium effect size of ≈ 0.5), would require 63 participants. Each participant was asked to rate 9 text excerpts, randomly divided between 4 conditions. Assuming an intraclass (i.e. between volunteer) correlation coefficient of 0.185 and a cluster size of 636/12 ≈ 31 gives a design effect of 2.48. The product of the design effect and the sample size for a non-repeating experiment is 2.48 x 63 ≈ 160, so we doubled that number to guarantee that if an effect was present that we would detect it, and we used the first 301 usable participants.

Data Available Online

Data is available online at: https://github.com/harry-witchel/CapitalizationMisspelling

Results

Differences between the Paragraphs

We ran an experiment in which we gathered data from 301 volunteers each with 9 experimental observations (2,709 ratings in total). The median trustworthiness ratings of each of the excerpts (E01 to E09) in the negative control condition (i.e. without any errors or incivility) are shown in the box and whisker plot in Figure TYK72z. On each box, the red line in the centre indicates the median, and the bottom and top edges of the box indicate the first and third quartiles, respectively. When the notches for two boxes do not overlap, this implies that the true medians do differ with 95% confidence. The whiskers extend to the most extreme data points not considered outliers, while outliers are shown using red plus signs; outliers are any points that fall more than 1.5 × the interquartile range away from the main box. Excerpt 01 generally elicits low ratings of trustworthiness (median = 27.5), while excerpts 08 and 09 generally elicit high ratings of trustworthiness (medians = 62 and 63, respectively). As illustrated by the non-overlapping notches, these two sets of excerpts elicit significantly different ratings of trustworthiness, which has been true in all our previous cohorts rating these excerpts for trustworthiness (data not shown). Excerpts 03, 04, 05 and 06 all elicited median ratings in the middle range from 45 to 55, while excerpt 02 is a transitional excerpt between E01 and the middle range and E07 is a transitional excerpt between the middle range and the most trustworthy excerpts.

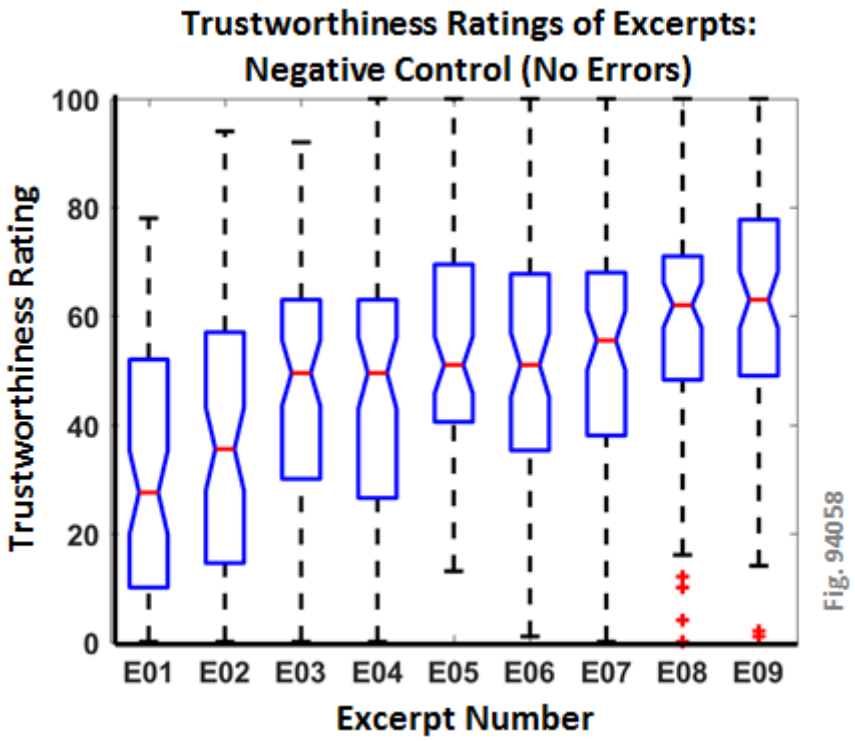

Figure TYK72z. Trustworthiness ratings of the different excerpts in the negative control (no errors) condition. For each box N = 75.

Cumulative Distributions Shifted Left by Errors

Figure RTL33A shows how the errors in writing mechanics and incivility led to changes in the cumulative probability distributions for each of the excerpts. As expected, for each excerpt, compared to the negative control with no errors (in black), all three alterations (inappropriate capitalization in blue, misspelling in magenta, and both errors together in red) led to a decrease in the ratings of trustworthiness (i.e. a leftward-upward shift in the curve). For most of the excerpts, at most points on the cumulative distribution curve, the combination of errors (both inappropriate capitalization plus misspelling) led to decreases in trustworthiness ratings (i.e. a shift left and up) compared to either of the single errors; that is, the lines for inappropriate capitalization only (blue) and misspelling only (magenta) fall between the black line (no errors) and the red line (both errors).

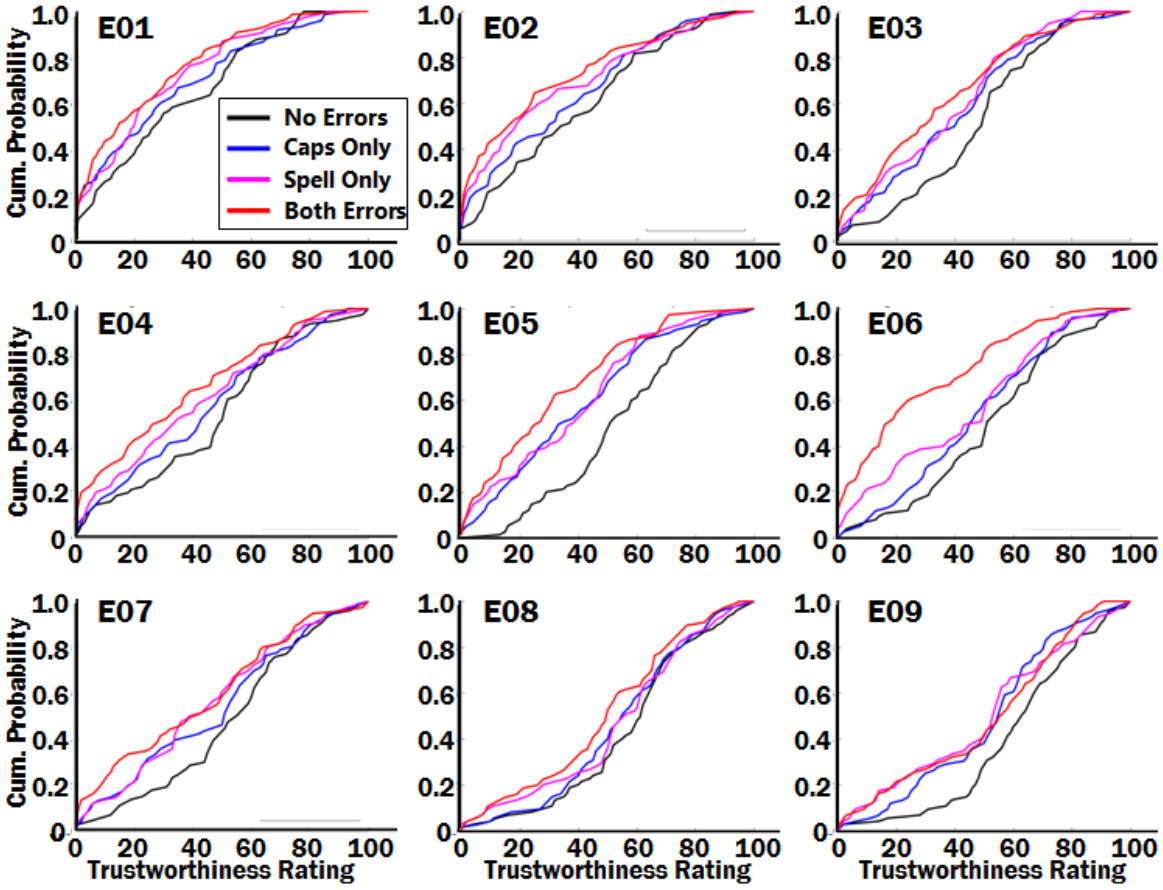

Figure RTL33A. Cumulative probability distribution plots for each excerpt (E01 to E09), comparing the alternative writing errors (coloured lines) to negative control (black).

Mixed Linear Effects Model with Four Conditions of Alteration

We tested this data in a mixed linear effects model (Model 1) where trustworthiness rating was the dependent variable, with two independent variables as fixed effects: text alteration (i.e. no errors (negative control), inappropriate capitalization only, misspelled only, and both errors) and excerpt (excerpts 01-09); the model also included a random effect to account for clustering of observations within volunteers. The reference group/condition for this model was E05 with no literacy errors. This model (and all of the following models) were calculated with robust standard errors [57 Williams 2000] to allow for the heteroskedasticity that is in this data set (see methods). The intracluster correlation (correlation within the individuals) coefficient estimate for Model 1 is 0.334 (95% CI: 0.287 to 0.384). The results of the LME model are shown in Table XiU77W.

Mixed Effects Regression		Number of Observations	2,709
Group Variable: Participant		Number of Groups	301
		Wald chi2 (11)	494.92
		Prob > Chi2	0.000
Log pseudolikelihood =	-12282.888		

(Std. Err. adjusted for 301 clusters in participants)

Categorical Variables	Coeff.	Robust Std. Err.	z	P >|z|	[95% Conf.	Interval]
Alteration						
Caps Only	-6.411	1.299	-4.93	0.000	-8.958	-3.864
Misspelled	-8.860	1.404	-6.31	0.000	-11.611	-6.109
Both Errors	-14.330	1.419	-10.10	0.000	-17.111	-11.550
Excerpt						
E01 (P2)	-11.936	1.708	-6.99	0.000	-15.284	-8.587
E02 (P7)	-8.287	1.580	-5.24	0.000	-11.384	-5.190
E03 (P3)	-0.703	1.564	-0.45	0.653	-3.769	2.363
E04 (P6)	0.034	1.540	0.02	0.983	-2.985	3.052
E06 (P9)	1.796	1.624	1.11	0.269	-1.388	4.979
E07 (P1)	6.187	1.714	3.61	0.000	2.827	9.547
E08 (P4)	14.749	1.529	9.65	0.000	11.753	17.745
E09 (P8)	14.224	1.462	9.73	0.000	11.359	17.090
constant	47.056	1.531	30.73	0.000	44.056	50.057

Random -	Effects	Parameters	Estimate	Robust Std. Err.	[95% Conf.	Interval]
Volunteer	Identity					
	var	(constant)	210.719	20.150	174.707	254.155
	var	(residual)	420.161	17.158	387.844	455.172

Mixed-effects regression — Number of obs = 2,709
Group variable: Volunteer — Number of groups = 301

Obs per group:
min = 9
avg = 9.0
max = 9

Wald chi2(11) = 494.92
Log pseudolikelihood = -12282.888 — Prob > chi2 = 0.0000

(Std. Err. adjusted for 301 clusters in Volunteer)

Rating	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
Style						
Caps Only	-6.411204	1.299488	-4.93	0.000	-8.958153	-3.864255
MSCap	-14.33029	1.418508	-10.10	0.000	-17.11052	-11.55006
Misspelt	-8.860079	1.403529	-6.31	0.000	-11.61095	-6.109213
Para						
para1	6.187413	1.714248	3.61	0.000	2.827548	9.547278
para2	-11.93578	1.708432	-6.99	0.000	-15.28424	-8.587315
para3	-.7029528	1.564132	-0.45	0.653	-3.768596	2.36269
para4	14.74907	1.528674	9.65	0.000	11.75293	17.74522
para6	.0336874	1.54016	0.02	0.983	-2.984972	3.052346
para7	-8.28701	1.580079	-5.24	0.000	-11.38391	-5.190112
para8	14.22442	1.4622	9.73	0.000	11.35856	17.09028
para9	1.795851	1.624215	1.11	0.269	-1.387553	4.979255
_cons	47.05641	1.531051	30.73	0.000	44.0556	50.05721

Random-effects Parameters	Estimate	Robust Std. Err.	[95% Conf.	Interval]
Volunteer: Identity				
var(_cons)	210.7194	20.14965	174.7068	254.1552
var(Residual)	420.1614	17.15765	387.8436	455.1721

Table XiU77W. Mixed Effects Model 1 for trustworthiness rating, with errors and excerpts as fixed effects and a random effect for the clustering of data by participant.

There is strong evidence against each of the null hypotheses that inappropriate capitalization only (Result 1), misspelled only (Result 2) and both errors (Result 3) do not affect trustworthiness ratings compared to the negative control group; stated positively, our data suggest that there is a statistically significant penalty to trustworthiness for inappropriate capitalization (Result 1), misspelling (Result 2), and for both errors together (Result 3). Note

that inappropriate capitalization only reduces trustworthiness ratings by -6.411 (95% CI: -8.958 to -3.864) and misspelled only reduces trustworthiness ratings by -8.860 (95% CI: -11.611 to -6.109). The effect on trustworthiness ratings of combining both inappropriate capitalization and misspellings together is -14.330 (95% CI: -17.111 to -11.550), which appears to be an additive effect.

Our further analysis aimed to test whether there was likely to be either an additive or integrative effect [37 Anderson 2013] of combining inappropriate capitalization and misspelling. Such an effect should lead to a significantly larger trustworthiness penalty when both error types are combined compared to either error individually. To test for this, an alternative specification of the LME model of the same data was formulated (Model 2). In Model 2 there were two separate independent Boolean variables for inappropriate capitalization only and misspelled only (as well as a categorical variable for the nine excerpts as in Model 1), so that the combination of both errors was not a single condition for the variable 'text alteration', but was considered as an interaction term of the capitalization and misspelling variables. In this model (Table JP92W3) the main effects and an interaction between them provide no evidence for an interaction effect between the two variables; that is, the main effects for inappropriate capitalization only and for misspelled only were as in the original Model 1 (Table XiU77W), while the coefficient for the interaction was not significantly different from zero. This supports the possible interpretation that the effects of the two error types are additive, rather than partially summative or than synergistic.

Dependent Variables	Coeff.	Robust Std. Err.	z	P >|z|	[95% Conf.	Interval]
Booleans						
Caps Only	-6.411	1.299	-4.93	0.000	-8.958	-3.864
Misspelled	-8.860	1.404	-6.31	0.000	-11.611	-6.109
Interaction	0.941	1.681	0.56	0.576	-2.355	4.237

Table JP92W3. Model 2: Alternatively specificed LME model with two Boolean dummy variables to test for any interaction between Misspelled only and Capitalization only. All unlisted values are identical to Table XiU77W above.

The output for Model 1 shows the following comparisons: (1) no errors versus capitalization only, (2) no errors versus misspelled only and (3) no errors versus both errors. Comparisons between (4) capitalization only versus misspelled only, (5) capitalization only versus both

errors and (6) misspelled only versus both errors can be made directly by changing the reference group in the model specification (Model 3 and Model 4):

Variable: Alteration	Coeff.	Robust Std. Err.	z	P >|z|	[95% Conf.	Interval]
Condition						
No Errors	6.411	1.299	4.93	0.000	3.864	8.958
Misspelled	-2.449	1.309	-1.87	0.061	-5.015	0.117
Both Errors	-7.919	1.204	-6.58	0.000	-10.279	-5.559

Table U7WU4d. Model 3. As per Model 1, but specifying inappropriate capitalization only as the reference condition.

Alteration	Coeff.	Robust Std. Err.	z	P >|z|	[95% Conf.	Interval]
Condition						
No Errors	8.860	1.404	6.31	0.000	6.109	11.611
Caps Only	2.449	1.309	1.87	0.061	-0.117	5.015
Both Errors	-5.470	1.202	-4.55	0.000	-7.827	-3.114

Table J7bV2C. Model 4: As per Model 1, but specifying misspelled only as the reference group.

Based on Models 3 and 4, (Result 4) there is weak evidence (P = 0.061) against the null hypothesis of no difference between the effects of capitalization only and misspelled only. Misspelled only leads to a larger trustworthiness penalty by -2.45 (95% CI: -5.02 to 0.12) compared to capitalization only, but this is not significantly different from 0 at 95% certainty level. There is strong evidence against the null hypothesis of there being no difference between the effects of capitalization only versus both errors combined (Result 5). Compared to capitalization only, the combination of both errors significantly reduces trustworthiness ratings by a further -7.92 (95% CI: -10.28 to -5.56). Likewise, there is strong evidence against the null hypothesis of there being no difference between the effects of misspelled only and the combination of both errors (Result 6). Compared to misspelling only, the combination of both errors leads to a further penalty to trustworthiness ratings of -5.47 (95% CI: -7.83 to -3.11).

Statistically Significant Differences between the Paragraphs

In Model 1, when compared to E05, the effect of the various excerpts' contents on trustworthiness ratings varies from -11.93 to 14.75 (a range of 26.68). This range is roughly twice as large as the effect of both errors (-14.33), suggesting that the errors in incivility and writing mechanics that we tested can together have an overall effect of nearly half of the effects of the content of the excerpts we tested.

Discussion

Original contributions

This study sought to quantitatively determine how two different errors of writing mechanics (contributing to incivility) combine to penalize subjective ratings of trustworthiness, in the medically relevant context of materials typical of an unmoderated online health forum. Using an LME model of a suitably powered study, we found that all three interventions (inappropriate capitalization, misspelling, and the combination of the two) were significantly different from the negative control (no added errors or incivility), which clearly answers RQ1. The data also show that (for these 70-100 word long excerpts about multiple sclerosis) the trustworthiness penalty for five instances of inappropriate capitalization was of a similar magnitude to the penalty for five instances of misspelling. Note that there was a trend for the penalty of misspellings to be larger, but as a generalized rule, the two are similar in magnitude and the precise difference will depend on exactly how many words and which words are capitalized or misspelled. This finding answers RQ2. Finally, with a combination of different Linear Mixed Effects Models, the data show that the combination of two different types of errors had a significantly greater trustworthiness penalty than either of the error types alone, and that the effect in this study was almost perfectly additive (RQ3); thus, the effects of the combination of errors was integrative [37 Anderson 2013] rather than a simplified heuristic such as take-the-best [45 Gigerenzer 1996], multiplicative, or affected by ceiling effects in this study. This begins to answer the question recently posed of how spelling interacts with other quality indicators on credibility [54 Sauls]. To the best of our knowledge, this is the first study that was specifically designed to test and quantify these kinds of specific additive effects on trustworthiness independently of the lay participants' insight.

We also showed that the stimulus excerpts that we designed are appropriate for studies that test trustworthiness of online health information independently of lay insight. While our

study did not preclude lay insight (i.e. participants might notice that some words were misspelled, etc.), the study was not dependent on such insight, which is useful for interrogating intuitive evaluations of information (i.e. cost-benefit analyses [6 Sun 2019]). In this study, the paragraphs engendered consistent effects on trustworthiness ratings (at least among this type of online psychology experiment cohort, see Figure TYK72z). Also, this is the first time that numerical values have been gathered for isolated effects of inappropriate capitalization.

The Results in Context

This study reaffirms an earlier observation that incivility decreases message credibility [35 Graf 2017, 36 Thorson 2010]. As suggested previously, inappropriate capitalization (shouting) is histrionic, and induces a strong effect of incivility on readers' subjective ratings [34 Gervais 2015]. In our study, the effect of shouting text showed a trend for eliciting a smaller trustworthiness penalty than the effect of similar quantity of misspelling. One could easily speculate about new experiments where we might change the quantities of literacy errors; our experiment used either five misspellings or five shouting phrases, but one could run experiments to titrate errors, for example to determine the relative effects of three misspellings or ten inappropriate capitalizations. Furthermore, the effects of text shouting may be moderated by whether the statements are controversial [35 Graf 2017]. We deliberately chose statements about multiple sclerosis that would be unfamiliar to the general public. This lack of familiarity makes the content/context not particularly emotional or controversial, likely weakening the effect of inappropriate capitalization.

The debate between how accessibility (i.e. cost and speed) versus information quality (i.e. accuracy and presumed benefit) quantitatively affect the use of (and search for) information continues [13 Woudstra 2016]. Categorical frameworks have long been proposed to provide a theoretical underpinning for the factors, grouping specific elements that influence perceptions of trustworthiness in a variety of ways. The most well-known two-category grouping of factors affecting how people respond to communication is Petty and Cacioppo's Elaboration Likelihood Model (ELM) for persuasion [53 Petty 1984], in which elements contributing to a central pathway (e.g. argument quality) are complemented by seemingly less rational elements that contribute to a peripheral pathway (e.g. website design) [38 Fogg 2003a, 20 Sbaffi 2017]. Another two-category persuasion model that has been used to explain online trustworthiness is Chaiken's dichotomy of heuristics versus systematic information [58 Chaiken 1980, 59 Sundar 2008]. In both the Heuristic-Systematic Model and

the ELM, diminishing motivation/involvement (or diminishing user dependency [42 Johnson 2015]) is associated with a switch from focusing on the effortful systematic evaluation of information (quality) to low-effort heuristics (accessibility).

In purchasing decisions, the type of product affects how strongly grammar and mechanics errors affect credibility [28 Ketron 2017]. In particular, 'experience goods' (non-technical items like body lotion that are used personally) are more affected by grammar and writing mechanics than 'search goods' (technical items like printers). The implication is that when objective signals about trustworthiness are absent, heuristics play a stronger role [58 Chaiken 1980, 53 Petty 1984]. In the current study, where laypersons made judgments about unfamiliar medical issues, we might expect to find a stronger response to misspelling and capitalization. A necessary future approach is to repeat this experiment with multiple sclerosis patients who would exhibit user-dependence when evaluating the statements [42 Johnson 2015].

Limitations

Our study had several important constraints. We deliberately tested laypersons to judge unfamiliar ideas about multiple sclerosis, and showed that literacy errors can have a strong effect. Nevertheless, this effect may be smaller in a cohort that is more dependent on knowing the information. In particular, if readers are dependent, then pre-existing points of view and feeling of 'homophily' [60 Wang 2008] will influence perceived credibility, in a way that would not influence the general public with less interest in statements for and about multiple sclerosis.

It is also notable that in the current study that the highest mean trustworthiness rating was 62.3 (0-100 scale), despite the statement being medically correct. Multiple factors may account for why this mean rating is not higher for trustworthiness. For example, participants in this psychology experiment saw the statements *in vacuo*, so that they could not verify the statement, check source credibility, or look for crowd consensus [8 Lederman 2014]. In real life situations, other aspects of trustworthiness may dominate, and among some individuals there may be ceiling and floor effects within this data set, given the very wide standard deviations.

Conclusions

Incivility and literacy competence are proposed factors in how lay web users assess trustworthiness of online health information. These results support Lederman et al.'s [8

Lederman 2014] theory of credibility assessment in online forums; the results also fit with Anderson's [37 Anderson 2013] description of integrative assessment of multiple cues. Here, we have shown that literacy competence errors have additive effects. Many other factors also contribute to trustworthiness, notably the argument quality of the content (logic), verification with other sources, reference credibility, and crowd consensus [8 Lederman 2014]. How these additional factors interplay with literacy competence will require further extensive research. A start would be to determine how persons with multiple sclerosis (with similar education levels to our current cohort) respond to these stimulus paragraphs, as those readers would understand the information in a more self-relevant context.

Acknowledgements

We gratefully acknowledge: Tom Ormerod and The School of Psychology at the University of Sussex for access to Qualtrics. We acknowledge Matthieu Raggett for IT set up. We also acknowledge the BSMS Independent Research Project programme for funding. Finally, we acknowledge the late Fayard Nicholas for the original idea on civility.

Competing Interests

BM owns and JR worked for a commercial enterprise (Dalton Maag) that creates typefaces and graphical designs for commercial clients. AN acts as an ad-hoc Scientific Consultant for Dalton Maag. This enterprise and those authors will not benefit directly or indirectly from data supporting incivility, misspelling, or shouting text.

HJW, GAT, CEIW, CIJ and HDC have no competing interests to declare.

Abbreviations

ANOVA – ANalysis Of VAriance statistical test.

ELM – the Elaboration Likelihood Model

df – degrees of freedom

F2F – face to face

LME – Linear Mixed Effects (statistical model)

SE – standard error

s.e.m. – standard error of the mean

SS – sum of the squares

Author Contributions

HDC, HJW, GAT, BM, AN & CEIW conceived the experiments. HJW, GAT, CEIW and JR conducted the experiments. HJW, CIP & GAT analysed the results. CIP performed the statistical analyses. HJW, GAT and HDC drafted the manuscript. All authors reviewed the manuscript.

References

1. Office for National Statistics UK (2019). Internet access – households and individuals, Great Britain: 2019: Internet access in Great Britain, including how many people have the internet, how they access it and what they use it for. [Accessed 2019-11-20]. https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2019

2. Chen YY, Li CM, Liang JC, Tsai CC. Health information obtained from the internet and changes in medical decision making: questionnaire development and cross-sectional survey. J Med Internet Res 2018;20(2):e47. PMID: 29434017

3. Hsieh RW, Chen L, Chen T, Liang J, Lin T, Chen Y, et al. The Association Between Internet Use and Ambulatory Care-Seeking Behaviors in Taiwan: A Cross-Sectional Study. J Med Internet Res 2016 Dec 07;18(12):e319. PMID: 27927606

4. The Health Union. Social media for health. 2016. https://health-union.com/wp-content/uploads/2016/06/Social-Media-for-Health.pdf. Archived at: http://www.webcitation.org/78aKBnoMO

5. Diviani N, van den Putte B, Giani S, van Weert JC. Low health literacy and evaluation of online health information: a systematic review of the literature. Journal of Medical Internet Research 2015;17(5):e112. PMID: 25953147

6. Sun Y, Zhang Y, Gwizdka J, Trace CB. Consumer Evaluation of the Quality of Online Health Information: Systematic Literature Review of Relevant Criteria and Indicators. Journal of Medical Internet Research 2019;21(5):e12522. PMID: 31045507

7. Fogg BJ, Soohoo C, Danielson DR, Marable L, Stanford J, & Tauber ER. How do users evaluate the credibility of Web sites?: a study with over 2,500 participants. Proceedings of the 2003 Conference on Designing for User Experiences; 2003; pp. 1-15. New York: ACM; 2003b, June. doi:10.1145/997078.997097

8. Lederman, R, Fan, H, Smith, S, & Chang, S. Who can you trust? Credibility assessment in online health forums. Health Policy and Technology 2014;3(1), 13-25. doi:10.1016/j.hlpt.2013.11.003

9. Stanford, J, Tauber, ER, Fogg, BJ, & Marable, L. 2002. Experts vs. online consumers: A comparative credibility study of health and finance Web sites. Consumer Web Watch. https://consumersunion.org/wp-content/uploads/2013/05/expert-vs-online-consumers.pdf.

Archived at: http://www.webcitation.org/78aNK0qXf

10. Brand-Gruwel S, Kammerer Y, Van Meeuwen L, Van Gog T. Source evaluation of domain experts and novices during Web search. Journal of Computer Assisted Learning 2017; 33(3), 234-251. doi:10.1111/jcal.12162

11. Sillence E, Briggs P, Harris P, Fishwick L. Going online for health advice: changes in usage and trust practices over the last five years. Interacting with Computers 2006 Nov; 28;19(3):397-406. doi:10.1016/j.intcom.2006.10.002

12. England CY, Nicholls AM. Advice available on the Internet for people with coeliac disease: an evaluation of the quality of websites. Journal of Human Nutrition and Dietetics 2004;17(6):547-559. PMID: 15546433

13. Woudstra L, van den Hooff B, Schouten A. The quality versus accessibility debate revisited: A contingency perspective on human information source selection. Journal of the Association for Information Science and Technology. 2016 Sep;67(9):2060-71. doi: 10.1002/asi.23536

14. Zimmermann, M., & Jucks, R. (2018). How experts' use of medical technical jargon in different types of online health forums affects perceived information credibility: Randomized experiment with laypersons. *Journal of Medical Internet Research*, 20(1), e30.

15. Kata A. Anti-vaccine activists, Web 2.0, and the postmodern paradigm–An overview of tactics and tropes used online by the anti-vaccination movement. Vaccine 2012 May; 28;30(25):3778-89. PMID: 22172504

16. Borzekowski DL, Schenk S, Wilson JL, Peebles R. e-Ana and e-Mia: A content analysis of pro–eating disorder web sites. American Journal of Public Health 2010 Aug;100(8):1526-34. PMID: 20558807

17.

Albuja, A. F., Sanchez, D. T., & Gaither, S. E. (2018). Fluid racial presentation: Perceptions of contextual "passing" among biracial people. Journal of Experimental Social Psychology, 77, 132-142.

Anderson, N. H. (2013). Unified psychology based on three laws of information integration. *Review of General Psychology*, *17*(2), 125–132. https://doi.org/10.1037/a0032921

BBC News Magazine. The cruellest of internet hoaxes. 2012. https://www.bbc.co.uk/news/magazine-18282277. Archived at: http://www.webcitation.org/78aSUcf7v t

Bennis WM, Katsikopoulos KV, Goldstein DG, Dieckmann A and Berg N (2012). Designed to Fit Minds. *In* Todd, P. M., Gigerenzer, G., & the ABC Research Group. (eds.) *Ecological rationality: Intelligence in the world*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195315448.003.0126

Castillo C, Mendoza M, Poblete B. Predicting information credibility in time-sensitive social media.

Internet Research, 23(5), 560-588 (2013). doi:10.1108/IntR-05-2012-0095

Chaiken S. Heuristic versus systematic information processing and the use of source versus message cues in persuasion. Journal of Personality and Social Psychology 1980;39(5):752-766. doi:10.1037/0022-3514.39.5.752

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. Proceedings of the National Academy of Sciences, 116(39), 19251-19257.

Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of internet e-surveys (cherries). Journal of Medical Internet Research 2004;6(3):e34. PMID: 15471760

Figueredo L, Varnhagen CK. Didn't you run the spell checker? Effects of type of spelling error and use of a spell checker on perceptions of the author. Reading Psychology 2005;26:441–458. doi:10.1080/02702710500400495

Flynn DJ, Nyhan B, Reifler J. The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. Political Psychology 2017;38, 127-150. doi:10.1111/pops.12394

Fogg BJ. Prominence-interpretation theory: Explaining how people assess credibility online. In CHI'03 Extended Abstracts on Human Factors in Computing Systems 2013 (pp. 722-723). New York: ACM; 2003a, April. doi:10.1145/765891.765951

Gervais BT. Incivility online: Affective and behavioral reactions to uncivil political posts in a web-based experiment. Journal of Information Technology & Politics 2015;12(2): 167-185. doi:10.1080/19331681.2014.997416

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, 103(4), 650-669.

Gigerenzer, G., & Goldstein, D. G. (2011). The recognition heuristic: A decade of research. *Judgment and Decision Making*, 6(1), 100-121.

Graf J, Erba J, Harn RW. The role of civility and anonymity on perceptions of online comments. Mass Communication and Society 2017;20(4):526-549. doi:10.1080/15205436.2016.1274763

Gray R, Fitch M, Davis C, Phillips C. Breast cancer and prostate cancer self‐help groups: reflections on differences. Psycho‐oncology 1996;5(2):137-142. doi.org/10.1002/(SICI)1099-1611(199606)5:2<137::AID-PON222>3.0.CO;2-E

Hobbs J, Kittler A, Fox S, Middleton B, Bates DW. Communicating health information to an alarmed public facing a threat such as a bioterrorist attack. Journal of Health Communication 2004 Jan;1:9(1):67-75. PMID: 14761834

Johnson F, Rowley J, Sbaffi L. Modelling trust formation in health information contexts. Journal of Information Science 2015 Aug;41(4):415-29. doi.org/10.1177/0165551515577914

Ketron S. Investigating the effect of quality of grammar and mechanics (QGAM) in online reviews: The mediating role of reviewer credibility. Journal of Business Research 2017;81:51-59. doi:

10.1016/j.jbusres.2017.08.008

Koriat, A. (2015). Metacognition: Decision-making processes in self-monitoring and self-regulation. *The Wiley Blackwell Handbook of Judgment and Decision Making*, 1, 356-379.

Kornfield R, Smith KC, Szczypka G, Vera L, Emery S. Earned Media and Public Engagement With CDC's" Tips From Former Smokers" Campaign: An Analysis of Online News and Blog Coverage. Journal of Medical Internet Research 2015;17(1):e12. PMID: 25604520

Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *American Journal of Occupational Therapy*, 45(3), 214-222.

Metzger MJ. Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research. Journal of the American Society for Information Science and Technology 2007;58(13), 2078-2091. doi:10.1002/asi.20672

Morin-Lessard E, McKelvie SJ. Does Writeing Rite Matter? Effects of Textual Errors on Personality Trait Attributions. Current Psychology 2017;1-12. doi:10.1007/s12144-017-9582-z

Oleson D (2013). Discovering Drug Side Effects with Crowdsourcing. data.World & Crowdflower. [Accessed 2019-11-23]. https://data.world/crowdflower/claritin-twitter/workspace/file?filename=1384367161_claritin_october_twitter_side_effects-1.csv

Oreg, S., & Bayazit, M. (2009). Prone to bias: Development of a bias taxonomy from an individual differences perspective. *Review of General Psychology*, 13(3), 175-193.

Orne, M. T. (2009). Demand characteristics and the concept of quasi-controls. *In* Robert Rosenthal and Ralph L. Rosnow (eds.), *Artifacts in Behavioral Research* (Oxford: Oxford University Press), pp. 110-137.

Pachur, T., & Bröder, A. (2013). Judgment: A cognitive processing perspective. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(6), 665–681. https://doi.org/10.1002/wcs.1259

Petty RE, Cacioppo JT. The effects of involvement on responses to argument quantity and quality: central and peripheral routes to persuasion. Journal of Personality and Social Psychology 1984;46(1):69-81. doi:10.1037/0022-3514.46.1.69

Pickard, V. Media failures in the age of Trump. The Political Economy of Communication 2017, 4(2), 118-122. http://www.polecom.org/index.php/polecom/article/view/74/264

Rieh SY, Danielson DR. Credibility: A multidisciplinary framework. Annual review of information science and technology 2007;41(1):307-64. doi.org/10.1002/aris.2007.1440410114

Sbaffi L, Rowley J. Trust and credibility in web-based health information: a review and agenda for future research. Journal of Medical Internet Research 2017;19(6), e218. PMID: 28630033

Schindler RM, Bickart B. Perceived helpfulness of online consumer reviews: The role of message content and style. Journal of Consumer Behaviour 2012;11(3):234-243. doi.org/10.1002/cb.1372

Sauls M. Perceived Credibility of Information on Internet Health Forums. Thesis. Clemson University 2018. https://tigerprints.clemson.edu/all_dissertations/2110. Archived at http://www.webcitation.org/78bQAEWKE

Seale C, Ziebland S, Charteris-Black J. Gender, cancer experience and internet use: a comparative keyword analysis of interviews and online cancer support groups. Social Science & Medicine 2016, 62(10), 2577-2590. PMID: 16361029

Sobieraj S, Berry JM. From incivility to outrage: Political discourse in blogs, talk radio, and cable news. Political Communication 2011;28(1):19-41. doi:10.1080/10584609.2010.542360

Sundar SS. The MAIN model: A heuristic approach to understanding technology effects on credibility. In Metzger MJ, Flanagin AJ. Digital Media, Youth, and Credibility. Cambridge, MA: The MIT Press; 2008:73-100. doi:10.1162/dmal.9780262562324.073

Thorson K, Vraga E, Ekdale, B. Credibility in context: How uncivil online commentary affects news credibility. Mass Communication and Society 2010;13(3):289-313. doi.org/10.1080/15205430903225571

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Vignovic JA, Thompson, LF. Computer-mediated cross-cultural collaboration: Attributing communication errors to the person versus the situation. Journal of Applied Psychology 2010;95(2):265-276. PMID: 20230068

Wanas N, El-Saban M, Ashour H, Ammar W. Automatic scoring of online discussion posts. In Proceedings of the 2nd ACM workshop on Information credibility on the web (pp. 19-26). ACM; 2008, October. doi:10.1145/1458527.1458534

Wang Z, Walther JB, Pingree S, Hawkins RP. Health information, credibility, homophily, and influence via the Internet: Web sites versus discussion groups. Health Communication 2008;23(4):358-368. PMID: 18702000

Weerkamp W, de Rijke M. Credibility-inspired ranking for blog post retrieval. Information Retrieval 2012;15(3-4):243-277. doi:10.1007/s10791-011-9182-8

Williams, R. L. (2000). A note on robust variance estimation for cluster‐correlated data. Biometrics, 56(2), 645-646.